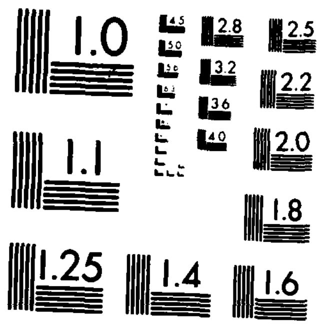
AD-A124 268   PROSPECTS AND PROBLEMS FOR A GENERAL MODELING METHODOLOGY(U) RAND CORP SANTA MONICA CA M DAVIS ET AL. JUN 82 RAND/N-1801-RC   1/1

UNCLASSIFIED   F/G 9/2   NL

END
DATE
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A124268

DTIC FILE COPY

A RAND NOTE

PROSPECTS AND PROBLEMS FOR A GENERAL MODELING METHODOLOGY

M. Davis, S. Rosenschein, N. Shapiro

June 1982

N-1801-RC

This publication was supported by The Rand Corporation as part of its program of public service.

Published by The Rand Corporation

A RAND NOTE

PROSPECTS AND PROBLEMS FOR A
GENERAL MODELING METHODOLOGY

M. Davis, S. Rosenschein,
N. Shapiro

June 1982

N-1801-RC

Rand
SANTA MONICA, CA. 90406

PREFACE

The study reported in this Note represents a first step toward alleviating some of the problems inherent in current computer modeling techniques. It suggests a form for a new modeling methodology free of many of the constraints of existing methods. This discussion should be of interest to researchers in modeling and simulation and to those responsible for the planning or sponsoring of such research. It may also be useful to persons who conduct, plan, or fund research in other activities that employ modeling.

Funding for this work was provided by The Rand Corporation from its own corporate research funds.

SUMMARY

Computer technology has yet to provide the modeling enterprise with adequate tools, with the results that

- o Modeling has become confused with simulation.
- o Modelers are artificially constrained to state-transition models.
- o Modelers cannot adequately deal with complex models.
- o Modelers often obtain numbers but not insight.
- o Quantitative and qualitative models cannot be gracefully combined.

Research that might alleviate these problems is discussed in this Note.

If such an ambitious program is to succeed, we must develop a language that can smoothly express quantitative and qualitative notions in the context of a stand-alone model. The Note describes one small step in this direction. We chose to model a little girl bouncing a ball. This situation involves both physical constraints and human decisionmaking, in a tightly intertwined manner. We have developed a model of this situation and an illustrative language for expressing that model. This exercise turned out to be very useful in demonstrating the importance of a stand-alone model and of the ability to smoothly combine qualitative and quantitative model elements.

The Note's program also depends on some improvement in automatic deduction. That issue is discussed.

ACKNOWLEDGMENTS

We appreciate the assistance and advice of Robert Anderson, Roy Goodman, Joyce Peterson, Hilary Putnam, and Ross Quinlan.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS vii

Section

I. INTRODUCTION 1

Definition and Development of Modeling 2

Simulators, Simulations, and Models 4

Inadequacies of Current Methods of Computer Simulation 7

Prospectus for a New Modeling Methodology 10

II. A LITTLE GIRL BOUNCES A BALL 12

III. AUTOMATED DEDUCTION 25

IV. SUMMARY AND CONCLUSION 29

I. INTRODUCTION

At least since the ancient Greeks, people have used models to understand and control aspects of reality. In recent years, models have been playing an ever more crucial role in our society. Models are used for prediction, for the vicarious or risk-free acquisition of experience, for the clarification of partially formulated ideas, and to communicate, control, store, and retrieve data. Unfortunately, the tools available for constructing the large, complex computer models required for such tasks have often proved inadequate for the demands placed on them. Although real-world situations usually have both qualitative and quantitative aspects, models constructed using currently available methods necessarily focus on one aspect and ignore the other.

In particular, models created to deal with complex social and biological processes are generally of a special restricted form: what are usually called **simulations**. Here, the restrictions under which existing methods have forced modelers to work have been compounded by terminological confusion. Since the only models in general use tend to be computer simulations, the habit has developed of using the terms "computer model" and "computer simulation" interchangeably, thus effectively obscuring the very possibility of breaking out of the confines of existing methods.

We believe that a new modeling methodology free of many of the constraints of existing methods is needed. This methodology should permit expression of both quantitative and qualitative elements in a single unified model that is conceptually independent of its computer implementation. The primary objective of this Note is to outline one form such a methodology could take and to indicate the kind of research needed to develop it into a useful system.

We shall present a discussion of some existing modeling methods and their limitations. Finally we shall study as a "toy" example the phenomenon of a little girl bouncing a ball, to show what will be involved in providing modeling that can be used to deal with both quantitative and qualitative relationships.

This Note barely scratches the surface with a preliminary exploration of possibilities. Serious development of the methodology we outline will be a large undertaking. But even modest practical success will be worthwhile. For the full scope of what we propose amounts to nothing less than automation of a significant part of the scientific and engineering enterprise.

DEFINITION AND DEVELOPMENT OF MODELING

We define a model as the more or less formal representation of an idealized aspect of reality. A model is obtained in two stages of abstraction from the real world. The first step is to simplify and idealize what is being studied, for the purpose of eliminating irrelevant complexities and permitting theoretical analysis. The second step is to state in precise (this usually means mathematical) language various key presumed facts about this idealization; these facts are the "axioms" of the model. Using the model then requires manipulating the axioms by appropriate formal techniques, which have the character of <deductions>, in order to obtain appropriate information.

The models that especially call for development of a new modeling methodology have traditional mathematical components, as well as additional organizational, socio-political, or human-behavioral components. We label these two classes of components, respectively, "quantitative," and "qualitative." The distinction between them is important in understanding the limitations of present computer models.

By "quantitative," we refer to the mathematical representation of relationships, from Newtonian mechanics to the contemporary models. Quantitative models tend to emphasize the numerical aspects of situations, use continuous rather than discrete abstractions, and be conceived of in terms of traditional mathematical notions such as algebraic expressions, integrals, and partial derivatives. Although quantitative models of some portions of reality are now quite sophisticated, they can not express logical relationships and non-numeric attributes of entities.

Computers have greatly amplified the power of preexisting models. Numerical answers can now be obtained quickly that in precomputer times would have required many man-years of computation, and even problems whose solution would once have been utterly unfeasible have now been brought into the range of what can be routinely accomplished. But this amounts to using computers only as a peripheral aid in the modeling process. The models themselves retain their "back-of-the-envelope" character, capable of succinct statement and of being understood as a single whole by those who use them. The steps in proceeding from the model to a computation are clearly defined and separate: First, relationships between quantities are deduced from the axioms of the model; next the variables are distinguished as independent and dependent, and an algorithm is specified for computing the values of the dependent variables given values of the independent variables as input (this step may well require the use of numerical analysis); finally, a computer program which implements this algorithm is written and run. This is in contrast to the use of computers as an integral part of the modeling process itself. Typical computer models--especially those developed in the biological and social sciences--are not succinctly expressible. They tend to exist only as (often quite large) computer programs. And they have made it possible to model processes that were beyond the scope of precomputer modeling methodology.

As already indicated, these computer models tend to be of a very special kind, **simulations**, leading to the unfortunate terminological confusion to which we have already alluded. We will argue that this limitation, although in large part forced on computer modelers by the nature of the software facilities available to them, has extremely undesirable consequences. The new methodology that we will outline would permit computers to manipulate models that have many of the virtues of back-of-the-envelope models but that are on a scale suitable for modeling complex social and biological processes.

SIMULATORS, SIMULATIONS, AND MODELS

Before the advent of computers, simulators were simply systems certain of whose properties were both easily determinable and linked to corresponding properties of a (usually larger and more complex) system about which information was desired. Examples are a model airplane in a wind tunnel as a simulator of actual aircraft in flight, a physical relief model of the Mississippi Valley (actually constructed for the U.S. Army Corps of Engineers) used as a simulator of water flow in the Mississippi Valley, or even an experimental group of individual human beings as a simulator of various kinds of social interaction. Each fixed use of such a simulator was a simulation, e.g., use of the wind tunnel with a certain airflow, pouring water in a particular way on the relief model of the Mississippi Valley.

Before computers were available, mathematical models had to be capable of direct manipulation by a user-scientist. This placed distinct (in fact very uncomfortable) bounds on the complexity of models. A model too complicated to be held all-at-once (or at least almost so) in the mind of the modeler could not be used, however well it mirrored the phenomenon under investigation. The coming of the large-scale digital computer opened a path to the use of really complex models. But so far the path has been narrow and constricted. To date, the software that has been made available by computer scientists and engineers to modelers is not designed for the general manipulation of models conceived as formal structures. The one option that has been available in the use of really complex models is the development from such models of computer programs which cause a computer to behave as a simulator of the phenomena being studied. Although such computer programs are often constructed in a rather ad hoc manner, we can describe what is conceptually involved: A model is developed (either directly or from a model already known) which gives a sequential description of the states of a system. Using a discrete representation of time, we can conceive of such a model as specifying the initial state of a system (i.e., at time 0) and giving precise rules for calculating the state of the system at time n + 1 from the state at time n. We will call such a model a **state-transition model**. A computer program that carries out the computations specified by such a state-transition model can then be regarded as a simulator. (More accurately, it is the computer executing the program which constitutes a simulator as we have been using the word, but we will not fuss over this pedantic distinction.) A computer simulation is obtained when a simulator is supplied with appropriate data (specification of the initial state as well as the values of parameters) and with necessary utility programs, e.g., report generators.

Widespread use of this kind of computer simulation has made possible practical solutions to problems that would otherwise have been intractable. But there has been insufficient fundamental analysis of what is really involved in this use of computers. Computer simulation has developed in such a natural and unselfconscious manner that not enough attention has been paid either to the implicit methodological presuppositions or to what is lost in restricting oneself to models that have such a very special character. In particular, the very fact that computer modelers have been laboring under this restriction is not only rarely mentioned, but has even been made difficult to perceive by the very language used; the terms "computer model" and "computer simulation" are often used more or less interchangeably.

INADEQUACIES OF CURRENT METHODS OF COMPUTER SIMULATION

Since we are proposing a new methodology for computer modeling, the development of which will require considerable resources, it is incumbent on us to show explicitly that the method of computer simulation currently employed has serious inadequacies.

Analyzing the behavior of a model is vital--indeed, it is usually the reason for the model's existence. If the model of a process is stated in mathematical notation, centuries of accumulated wisdom can be brought to bear in its analysis. The symbolic formula can be differentiated and, perhaps, the sign of a quantity can be determined over an interval, using purely analytic--as opposed to numeric--techniques. That information might lead to important simplifications and insights, without our performing a single numerical calculation. In fact, we believe that information can be communicated more effectively to people through such general statements as "X increases when Y/Z decreases" than through computational results that imply that relationship. Such statements are perhaps best generated by applying analytic tools directly to the symbolic representation of a model. If the only formal representation of a model is a FORTRAN program, the power of such analytic techniques is lost. The code can only be executed, often resulting in numbers, not insight.

It is the underlying model that people want to operate on, to understand, to communicate to one another, even to alter. In practice, it is often the case that the only complete description available of a computer simulator is the computer program itself. The designers will not have troubled to specify explicitly the underlying state-transition model on which they based the program. As a result, a host of assumptions of very different degrees of importance and reliability are buried in the program.

These assumptions are almost impossible to retrieve, especially if it has been some time since the program was written, and therefore they are almost impossible to assess or change. They are also commingled in ways that make it very difficult to know the effect of each on some important conclusion being drawn from the simulation. Parts of the code may represent assumptions that are regarded as accurate and well-founded. Other parts may represent assumptions that are regarded skeptically or as being only approximately correct. Others may be derived by using numerical analysis with an appropriate error analysis.

One of the dangers implicit in the use of computer simulators is that the modeler may be forced to represent an assumption that is well believed but is not equivalent to a state-transition assertion (e.g., an inequality instead of an equation) by a code which expresses a far stronger relationship than experts in the field would be willing to assert as reliably known. Thus, a computer simulation might give results having a high degree of precision that is in no way justified by what is really known about the phenomenon being simulated.

These considerations make clear that it may be difficult to infer from the results of a computer simulation how sensitive those results are to the particular values of parameters that may have been used. When there are many such parameters, vitally important sensitivity analyses either cannot be carried out at all or can only be carried out in a limited and/or prohibitively expensive manner. This difficulty in conducting sensitivity analysis can be crucially significant when simulations are used in making important decisions.

Even if we ignore all the above considerations, computer simulations have serious limitations as models. State-transition models, in principle, can do only one kind of analysis, can answer only one kind of question: If a system has a given (total) state at time 0, what is its state at time n? This question may or may not correspond to what a user really needs to know. Furthermore, a state-transition model cannot perform any analysis unless it has total information about a state, whether or not this is needed for the problem at hand. A user cannot ask a computer simulation, "What initial state will produce a certain desired state at time n?" or "Given that at time n two key parameters have certain values (i.e., one knows part but not all of the state at time n), what can one say about these parameters at time n + 1?" A user can never derive a general statement such as "if x > y at time 0, then x > y at all subsequent times," using a computer simulator.

We believe that an intangible but critically important consequence of these limitations is the effect they have had on the mindset of computer scientists and users. Information about the domain being modeled must be expressed in the state-transition format no matter what violence this procrustean treatment does to our knowledge of the domain. Consequently, we have conditioned ourselves to model domains in terms of what the simulator can do rather than what we know or need to know about the domain's possibly rich mix of quantitative and qualitative elements.

PROSPECTUS FOR A NEW MODELING METHODOLOGY

Developing New Tools

Considering the growing importance of modeling, what do we propose should be done about the inadequacies described above? As a minimum goal, we would hope that these ideas will become well understood by computer modelers, who will be led to construct computer simulations in ways that meet some of our objections. In particular, this would call for explicit statement of the underlying state-transition model in mathematical language and a conscious separation of this model from the computer program (i.e., simulator) that implements it. The program itself should be written in a modular fashion so that it will be clear at each point which parts of the program correspond to which assertions of the underlying state-transition model. With this minimum policy one would hope to do for computer simulation what the structured programming movement has tried to do for computer programming.

But this minimum policy falls far short of what we feel to be necessary. If computer technology is to begin to realize its potential as a tool in the hands of modelers, information scientists will have to provide an entirely new comprehensive modeling methodology. This methodology should make it possible for modelers to express the most diverse relationships--qualitative as well as quantitative--in computer manipulable form. These relationships might take the form of algebraic or differential equations, inequalities, logical interrelationships among qualitative assertions, or even relationships among structures of a kind not even envisioned by the designers of the underlying modeling system (e.g., Feynman diagrams). Recognizing that the step from state at time n to state at time n + 1 in state-transition models is really a form (albeit an extremely limited form) of deduction, we must permit the computer modeler to use the computer itself in making necessary deductions of whatever sort from the basic relationships of the model.

We are thus envisioning a computer modeling facility in which a user would be provided with a comprehensive but comfortable language, extensible as necessary and capable of expressing the most diverse relationships, together with efficient automated and semi-automated deduction capabilities. The development of such a facility is of course an enormously ambitious project that can be expected to require sustained efforts over a considerable period. Nevertheless, if we are to make proper use of the rapidly developing computer technology as a basis for modeling, such an effort should be begun.

II. A LITTLE GIRL BOUNCES A BALL

In developing the kind of comprehensive modeling facility we have been envisioning, various crucial design decisions will need to be made. Such decisions can be made in an intelligent manner only after preliminary efforts to construct models of phenomena similar to those which the system being developed is expected to model. Ideally such a facility should be able to model phenomena involving complex combinations of physical, biological, psychological, and social factors. Some of the axioms, particularly those dealing with the physical aspects, can be expected to be "quantitative," that is, to take the form of equalities (numerical or matrix), differential, difference, or integral equations, and inequalities. In addition, we must expect axioms of a more "qualitative" nature, which, for example, express logical relationships. Some axioms will have both quantitative and qualitative features.

As a first approach to modeling phenomena that involve both people and physical systems, we have chosen to attempt to model a little girl bouncing a ball. Our model will thus have to deal with the physics of the bouncing ball as well as the behavior of the little girl. We are not going to attempt a complete model of this phenomenon. Rather, we shall exhibit and discuss some appropriate axioms and their consequences. This will be enough to indicate how a rich and complicated phenomenon having quantitative as well as qualitative or even psychological aspects can be modeled. We will not attempt to delimit an appropriate formal language in which our model could be written. However, some necessary features of such a language will emerge from the mathematical form of our axioms.

Consider the behavior of a baseball outfielder who observes a ball being hit by a batter and runs to successfully position himself underneath the ball. An appropriate model of the ball's behavior is that it travels in a parabola, its height at time t (measured from the instant the bat strikes the ball) being given by: $h = At + Bt^2$ for suitable constants A, B. What might an appropriate model of the ballplayer's behavior be? If we think in terms of a "real" person and try to imagine what is "really" happening in his mind and body, the obstacles (in the present stage of human knowledge about such matters) are surely insurmountable. What if instead we are saticfied with a model that successfully predicts his behavior? Then we may introduce a notion of "belief," as in "The ballplayer believes," and seek appropriate axioms governing this notion. One such axiom might take the form:

> The ballplayer believes the batted ball travels in the parabola $h = At + Bt^2$.

It is both difficult and important to learn to work with such assertions in a nonanthropomorphic manner. One may perfectly well adopt such an axiom knowing that the particular ballplayer being modeled never could pass algebra in school and knows nothing of parabolas.

The point of view being suggested has some points of contact with issues in the philosophy of science. According to some schools of thought, the correctness of a scientific theory (which is not very different from what we have been calling a model) is entirely a question of its ability to correctly predict what is observed. Others insist that more is required. Fortunately, we can remain entirely neutral on these issues. For us it is enough that the usefulness of models for our purposes depends only on their ability to yield correct predictions.

In our model, time will be modeled as a real positive parameter t. We write T for the set of positive real numbers. Various physical objects (including the little girl) will be modeled as sets of small, disjoint parts called particles. In particular, GIRL is to be the set of particles making up the little girl. We shall use the letters p, q, r for particles. The ball will be modeled as a particle. For any particle p (which may or may not be a part of the girl) $cg_t(p)$ is to be the center of gravity of p, a point of E^3 (three-dimensional Euclidean space), at time t. Similarly, $cg_t(\text{Ball})$ is to be the center of gravity of the ball at time t. We will write $\Box_t\ \alpha$ to mean that the girl "believes" or "knows" α at time t. (More will be said about this later.)

Some of our axioms will record appropriate information about the physical behavior of the girl's parts. For example:

$$(\forall t \in T)\ (\forall p,q \in \text{GIRL})\ [cg_t(p) = cg_t(q) \rightarrow p = q]\ .$$

(This axiom expresses the fact that no two parts of the girl's body can be in the same place at the same time.) We think of E^3 as a real inner-product space, so that for any x, y $\in E^3$, angle (x,y), where $0 \leq$ angle $(x,y) < \pi$, is well-defined (e.g., as the angle whose cosine is the inner product of x and y divided by $|x|\bullet|y|$). Another pair of axioms give restrictions on the girl's ability as a contortionist. Namely,

$$(\forall p,q \in \mathrm{GIRL})\left[\left|\frac{d}{dt}\,|cg_t(p) - cg_t(q)|\right| < A(p,q)\right]$$

$$(\forall p,q,r \in \mathrm{GIRL})$$

$$\left[\left|\frac{d}{dt}\,\mathrm{angle}\,(cg_t\,(p)) - (cg_t\,(r)),\ cg_t\,(q) - (cg_t\,(r))\right| < B(p,q,r)\right] .$$

Here the bounds A,B will change as the girl's abilities to manipulate her body improve.

Axioms can be given which govern the motion of the ball. For example, the following expresses the ball's "law of motion":

$$(\exists F \subseteq T)\left[F \text{ is finite } \& \ (\forall t \in T-F)\ \frac{d^2}{dt^2}\,cg_t(\mathrm{Ball}) = g(t)\ \right] .$$

Here F is to allow for the instants when the ball comes in contact with an obstacle and "bounces" or stops; g(t) gives the force of gravity and (at least if we confine ourselves to the surface of the earth) could be taken to be constant. Additional axioms can be written to express the elementary geometry and physics of the ball's "bouncing." These axioms will involve a coefficient of elasticity λ.

Now we come to axioms involving $\Box_t$. From a logical point of view, $\Box_t$ satisfies formal rules like those satisfied by so-called "modal" operators, and indeed this consideration has suggested the notation we are using. Our intuitive idea is that the girl herself works with a model of the bouncing ball. The interpretation we suggest for "$\Box_t\ \alpha$"

is: <u>The sentence α is part of the girl's model at time</u> t. It turns out to be surprisingly difficult to avoid the mistake of anthropomorphizing the "little girl." For naturally, when we write axioms of the form $\Box_t \alpha$, we in no way intend to indicate any belief on our part that the flesh-and-blood little girl truly believes α. What we are claiming is only that the <u>behavior</u> of the flesh-and-blood little girl will be predicted by a model of her in which she believes α. To help in avoiding this confusion we will write the word GIRL in capitals to indicate our model of the girl, whereas we will reserve lowercase for the flesh-and-blood little girl. Thus when $\Box_t \alpha$ is in our model, we may say: The GIRL believes (or knows) α (at time t). As an example, the

perception axiom:

$$(\forall t, t' \in T)(\forall p,q)\ (\forall x,y,z \in E^3)$$
$$\{\Box_{t'}\ [cg_t(p) = x \ \&\ cg_t(q) = y \ \&\ cg_t(eye) = z]$$
$$\rightarrow |\text{angle } cg_t(p)-cg_t(eye),\ cg_t(q)-cg_t(eye) \ - \text{angle } (x-z,y-z)|$$
$$\leq C(t,t')\} .$$

Roughly speaking: If the GIRL believes that at some past (or future) time a pair of particles and her own eye will be at certain places, then at that past (or future) time, the location of the particle bears a definite relation to the belief.

Another set of axioms bear on the GIRL's ability in logic. Let Γ be a suitable set of axioms for the predicate calculus requiring only modus ponens as a rule of inference. Then we assume **completeness axioms:**

$(\forall t \in T) \Box_t \alpha$, for all $\alpha \in \Gamma$.

$(\forall t \in T)(\forall \alpha,\beta)[(\Box_t \alpha \;\&\; \Box_t (\alpha \rightarrow \beta)) \rightarrow \Box_t \beta]$.

Together these axioms imply that the GIRL can carry out any logical deductions whatever. (Needless to say, it is only the GIRL to whom we are imputing such virtuosity, not the little girl.) It may also be helpful to assume that the GIRL preserves a high degree of self-consciousness:

$(\forall t,t' \in T)(\forall \alpha)[(t' \geq t \;\&\; \Box_t \alpha) \rightarrow \Box_{t'} \Box_t \alpha]$.

What does the GIRL know about the ball? We assume that the GIRL knows the axioms governing the ball's motion. In particular,

$$(\forall t' \in T)\ [\Box_{t'}\ (\exists F \subseteq T)\ F \text{ is finite } \&$$

$$(\forall t \in T\text{-}F)\ \frac{d^2}{dt^2}\, cg_t(\text{Ball}) = g(t)\ ]\ .$$

That is, the GIRL knows the ball's law of motion. However, note that the gravitation function G is permitted to depend on t'. This permits the GIRL to "learn" better values for G. And in fact, as she becomes more skilled at ball bouncing, the value of G(t,t') will tend toward the true value g(t). Also, we are assuming that the GIRL knows axioms about bouncing. In this case, the value λ of the coefficient of elasticity is a function of t' and may change with the GIRL's experience.

An extremely important assumption is the **autonomy axiom:**

$(\forall t \in T)(\forall p \in \text{GIRL})(\forall x \in E^3)[\Box_t(cg_t(p) = x) \rightarrow cg_t(p) = x]$.

Thus, if the GIRL knows at a given time that some part of her body is at a given place, then it really is. The key role of the autonomy axiom is to relate what the GIRL believes to what actually holds. As a consequence of the autonomy and completeness axioms we can show that the GIRL's beliefs at any time are consistent. For, suppose that for some sentence α, $\Box_t\ \alpha$ and $\Box_t \sim \alpha$. By the completeness axioms, we have first

$$\Box_t\ (\alpha \rightarrow (\sim \alpha \rightarrow \beta))$$

for any sentence β whatever (since $\alpha \rightarrow (\sim \alpha \rightarrow \beta)$ is a tautology and hence can be deduced from Γ using modus ponens), and then (using modus ponens twice), $\Box_t\ \beta$. Since β is arbitrary, we may choose it to have the form: $cg_t(p) \neq x$ where in fact $cg_t(p) \neq x$. But this contradicts the autonomy axiom. Note also that the autonomy axiom actually prevents the GIRL from having certain beliefs. She cannot believe that any parts of her body will have positions other than their actual positions.

A final axiom which is useful for technical reasons is the **self-awareness axiom:**

$$(\forall t \in T)(\forall p,q \in \text{GIRL})\ (\forall x,y \in E^3)$$

$$[\Box_t(cg_t(p) = x \vee cg_t(q) = y) \rightarrow \Box_t cg_t(p) = x \vee \Box_t cg_t(q) = y].$$

This axiom expresses the GIRL's inability to know that one or another part of her body is in some particular place without knowing which part and where it is.

So far our axioms have imputed to the GIRL only "rational" beliefs. But it should not be assumed that our modeling framework is limited to such. The GIRL may believe that good ball bouncing requires nightly recitation of prayers, or that the ball's intersecting pavement cracks must be avoided. There is no difficulty in incorporating such beliefs by adding appropriate axioms.

For an illuminating deduction from our axioms, let us suppose that our GIRL has tired of mere ball bouncing and has become devoted to pinball machines. No change in our axioms is required; it is only necessary to think of "ball" as representing one of the balls used by the pinball machine. We suppose that the machine is equipped with a pair of "flippers," one manipulated by a button activated by a finger of the GIRL's left hand, the other by a button activated by a finger on the GIRL's right hand. If she succeeds in hitting the ball with one of the flippers, the ball remains in play. Otherwise it passes out of play. We let t be the instant when the flipper could be activated and let $t' > t$ be a subsequent instant such that the equation $cg_{t'}(\text{Ball}) = x_0$ signifies that the ball is now passing out of play. Let ℓf, rf be the fingers in question. (Thus, $\ell f, rf \in$ GIRL.) The left (right) button being activated can then be rendered $cg_t(\ell f) = z_\ell$ ($cg_t(rf) = z_r$). The relationship we have been discussing can be represented by the following equivalence (which we will call γ):

$$(1)\quad cg_{t'}(\text{Ball}) \neq x_0 \leftrightarrow (cg_t(\ell f) = z_\ell \vee cg_t(rf) = z_t) .$$

We suppose that (1) is not only true, but that the GIRL knows (1) at time t:

$$(2)\qquad\qquad \Box_t\, \gamma .$$

Furthermore, we suppose that the GIRL knows at time t that the ball will remain in play:

(3) $\quad \Box_t \; (cg_{t'}(\mathrm{Ball}) \neq x_0)$.

We proceed to show that the statement (3) of the GIRL's belief has consequences for the "external" world. Namely, we claim that our axioms together with (1), (2), (3) imply:

(4) $\quad cg_{t'}(\mathrm{Ball}) \neq x_0$.

The proof is quite simple. By the completeness axioms, (2) and (3) imply:

$$\Box_t \; (cg_t(\ell f) = z_\ell \vee cg_t(rf) = z_r) \; .$$

By the self-awareness and completeness axioms:

(5) $\quad \Box_t cg_t(\ell f) = z_\ell \vee \Box_t cg_t(rf) = z_r$.

Now suppose that (4) is false:

$$cg_{t'}(\mathrm{Ball}) = x_0 \; .$$

Then, by (1),

$$\sim(cg_t(\ell f) = z_\ell \vee cg_t(rf) = z_r) \; ,$$

i.e.,

$$cg_t(\ell f) \neq z_\ell \;\&\; cg_t(rf) \neq z_r \; .$$

So we have separately:

$$cg_t(\ell f) \neq z_\ell$$

$$cg_t(rf) \neq z_r \; .$$

By the autonomy axiom, these imply:

$$\sim\Box_t(cg_t(\ell t) = z_\ell)$$

$$\sim\Box_t(cg_t(rf) = z_r) \ .$$

But these last contradict (5).

What is interesting about this deduction is that <u>we</u> have no idea whether it was the left or the right flipper which was in fact activated. (Of course, by the self-awareness axiom, the GIRL does know, but <u>we</u> do not.) <u>This kind of indeterminate disjunctive inference is in principle impossible for a state-transition model</u>.

We now shift our example once again, this time to the GIRL learning to serve the ball in a tennis game. We can think of the tactics of tennis service as involving a tension between the desire to hit the ball as hard as possible (so as to maximize the probability that the opponent will be unable to return the ball) and the desire to avoid hitting the ball so hard that it will go out of bounds. (In this simplified version we are ignoring complicating factors such as spin.) Traditional modeling techniques would suggest the use of probability theory. An axiom might, for example, state that the ball is hit in such a manner that the probability that it goes out of bounds is 20 percent. Now there is nothing at all objectionable in the use of probability theory, and there is no special difficulty in incorporating probabilistic considerations into the kinds of models we have been discussing. However, it is very interesting that by using our "belief" operator, we can, without using probability, easily express the GIRL's "tactic" of trying to hit the ball so as not to go out of bounds, but otherwise, to hit the ball as hard as possible. We shall use two axioms: the first is to the effect that the GIRL will actually carry out any persistent future intention concerning parts of her body; the second that she intends (or plans) to hit the ball so as to maximize its velocity subject to the constraint that she not "know" that she is hitting it out of bounds.

Let $p = (p_1, p_2, \ldots, p_n)$ be a vector of parts of the GIRL's body, and let $\underline{x} = (x_1, x_2, \ldots, x_n)$ be a corresponding vector of points of E^3. We write:

$$S_{\underline{p},\underline{x}}(t) = \bigwedge_{i=1}^{n} (cg_t(p_i) = x_i) .$$

Note that by the autonomy and completeness axioms, we have for any $\underline{p}$, $\underline{x}$:

$$\Box_t \, S_{\underline{p},\underline{x}}(t) \rightarrow S_{\underline{p},\underline{x}}(t) .$$

The **plan fulfillment axiom** states:

$$(\forall t)(\exists t_0 < t)(\forall \underline{p})(\forall \underline{x})[(\forall \tau)(t_0 \leq \tau < t \rightarrow \Box_\tau \, S_{\underline{p},\underline{x}}(t)) \rightarrow \Box_t \, S_{\underline{p},\underline{x}}(t)] .$$

Next, let $\underline{p}^0 = (p_1^0, \ldots, p_n^0)$ be the particular array of parts of the GIRL's body whose positions determine how the ball will be hit by the racquet. For each vector $\underline{x} = (x_1, \ldots, x_n)$ let $g_t(\underline{x})$ be the velocity imparted to the ball (possibly 0) under the configuration

$$cg_t(p_i^0) = x_i \; , \; i = 1,2,\ldots,n .$$

Let ω stand for the statement "the ball is hit out of bounds," and let $M(\underline{x},t,\tau)$ abbreviate the statement:

$$\Box_\tau \left(S_{\substack{0\\ \underline{p},\underline{x}}}(t) \to \omega\right) .$$

Thus, $M(\underline{x},t,\tau)$ asserts that at time τ, the GIRL knows that if at time t her parts are in the configuration corresponding to $\underline{x}$, then the ball will be hit out of bounds. The **plan to hit-the-ball hard axiom** then states:

$$(\exists t_0 < t)(\forall \tau)(t_0 \leq \tau < t \to$$

$$\Box_\tau (\exists \underline{x}) \{S_{\substack{0\\ \underline{p},\underline{x}}}(t) \; \& \sim M(\underline{x},t,\tau)$$

$$\& \; (\;\underline{y})[\sim M(\underline{y},t,\tau) \to g_t(\underline{x}) \geq g_t(\underline{y})]\}) .$$

These methods can be used to accommodate all manner of planning and tactical situations. The problem of the "contrary-to-fact" implications such as "if I hit the ball thus-and-so, then it will go out of bounds" has been quite simply finessed.

The plan fulfillment axiom sheds interesting light on the phenomenon of "follow-through" in such sports as tennis, golf, and baseball. The follow-through is the trajectory of the racquet, club, or bat <u>after</u> the ball has been struck. Players of these sports seek to achieve a "good" follow-through. On the face of things, this is a bit paradoxical: How can events which take place after the ball has been hit have any effect on the ball's behavior? The answer is that what one strives to learn is how to <u>plan</u> for an effective follow-through. Since this planning takes place (in part) before the ball is hit, it can perfectly well influence how it is hit. This is an example of an interesting insight obtained by reflecting on our model which could hardly have been obtained using state-transition models.

Straightforward extensions of our methods should enable us to deal with interactions among more than one individual. We can write, e.g., $\boxed{i}_t\, \alpha$ to mean that the i-th individual knows or believes α at time t. Then we can easily express such statements as "Mary serves to the rear left court because she thinks that Sue expects her to move forward and to the right."

We hope that despite the fragmentary nature of our example, we have made clear both the possibilities inherent in the new modeling methodology we propose and the large amount of work that remains before they can be realized.

III. AUTOMATED DEDUCTION

The use of any model to obtain conclusions about the relevant subject matter involves deduction. Indeed, any technique for drawing conclusions from the axioms of a model can be regarded as that of carrying out a deduction. In particular, this is true of the traditional state-transition models. But then the deductions are of a severely limited form. Unfortunately, the full use of a general modeling methodology of the kind we are discussing would require far more versatile deduction technology than is available today. This is a field in which there is a great deal of ongoing research. We have nothing new to contribute, and therefore we content ourselves with a brief survey of what is available.

The really spectacular contribution of computer technology is of course in the area of numerical computation. Naturally, a usable computer modeling technology should make it easy for a user to make numerical calculations as needed in working with a model. What we wish to emphasize here is that such numerical calculations are themselves deductions. When we solve the equation $x^2 = 2$ to obtain $x = 1.414$ "correct to 3 places," we are deducing from the premise $x^2 = 2$, the conclusion $1.4135 \leq x < 1.4145$.

Computer facilities also exist for efficient algebraic manipulation. Packages such as MACSYMA and REDUCE can be used quite effectively to work with algebraic and the elementary transcendental functions; they can also be used to carry out formal differentiation and integration. The use of such a capability would be very useful in working with a model whose axioms consist of equations (including differential equations) and inequalities.

However, a general modeling technology cannot be restricted to axioms of this special form. Hence, we must seek to make available facilities for deductive reasoning of the *most general* kind. To begin with, this includes the propositional calculus, i.e., Boolean reasoning. This encompasses the logical relations among sentences which hold simply by virtue of their decomposition using the Boolean connectives $\sim$, &, V, $\rightarrow$. Now, the problem of determining whether or not a given conclusion follows from given premises in this way is known to belong to the class of combinatorial problems called NP-complete, and the problems in this class are thought to be computationally intractable. Nevertheless, there are algorithms for testing such inferences which work very well in practice. (The "worst" cases on which these algorithms are intractable rarely show up in practice.) So there will be no practical difficulty in making a Boolean deduction capability available to modelers.

Logical deduction in mathematics requires using the quantifiers $\forall$ and $\exists$ as well as the Boolean connectives, i.e., the full predicate calculus. For the predicate calculus, the problem of determining whether a given conclusion follows from given premises is known to be unsolvable, i.e., no algorithm exists. In this difficult situation, three approaches have been used, but so far none of them has been really satisfactory. One approach is to find important special cases for which algorithms do exist. Although there has been a great deal of interesting work by logicians on this approach, it is of little practical value. This is because on the one hand the cases included are too restrictive, and on the other, the algorithms are not very efficient. Another approach, which has generated a large literature, is known as mechanical theorem-proving. This approach uses algorithms that search for proofs in predicate calculus, terminating only if a proof is found. The search space can be drastically limited in various ways, but except for problems of very special forms, efforts to prove theorems in this way are defeated by a combinatorial explosion. The final approach is to rely on man-machine interaction, that is, to use so-called "proof checkers." The user is expected to supply the steps in a deduction, and the machine verifies their correctness. There are a number of proof checkers in operation. They tend to be almost painful to use because of the small deductive steps to which they are limited. An improved variant would use a modest mechanical theorem-proving capability to enable a proof checker to take more substantial steps. Systems of this kind are under development and preliminary indications are hopeful.

The deductive facility available to a modeler should allow all of the capabilities we have been discussing, with appropriate articulation among the various kinds of deduction. A user should be able to call, as needed, for numerical calculation, algebraic manipulation, propositional calculus testing, or predicate calculus theorem-proving/proof checking. A single proof being developed may well contain steps of all of these kinds, and the general system must be capable of dealing with these steps as part of an integrated proof. There are serious technical problems to be solved in order to carry out such an implementation. But this is a problem to be faced not only in connection with a general-purpose modeling technology, but also for any of the many purposes for which a comprehensive, versatile automated deduction capability is important.

Models like our little GIRL which involve a knowledge operator $\Box_t$ require an implementation of the completeness axioms. In the presence of a predicate calculus theorem-prover and/or proof checker, this is easily supplied. One needs only two rules of inference:

$$\frac{\Box_t\ (\alpha \rightarrow \beta) \qquad \Box_t\ \alpha}{\Box_t\ \beta}$$

and

$$\frac{\gamma}{\Box_t\ \gamma}$$

where γ has been proved <u>with no use of axioms</u>. In addition to closure under modus ponens, it may be desirable (although it is not absolutely essential) to add closure rules under other rules of inference of propositional calculus, e.g.,

$$\frac{\Box_t\ \alpha \qquad \Box_t\ \beta}{\Box_t\ (\alpha \ \&\ \beta)}$$

Finally, the deductive capability supplied to a modeler should be user-extensible. This is in order to allow for formal deductive methods specific to some model which have not (yet) been incorporated into existing rigorously justified mathematics. Examples are Feynman diagrams in quantum field theory and, earlier, the Heaviside operational calculus.

IV. SUMMARY AND CONCLUSION

Existing computer models are almost all of one very special type--state-transition models. Implementations tend to merge model and implementation in ways that produce methodological and practical difficulties.

We propose a new methodology by which modelers would be provided with a general-purpose language for specifying models of arbitrary character as well as automated and semi-automated deductive capabilities for their formal manipulation. Even without computer-aided deduction, such a language could be useful simply for the precise specification of models, just as very high-level programming languages can be useful for specifying algorithms, even in the absence of an implementation. With the current limited state of the art of automated deduction, we could still hope to offer the user a number of special-purpose deductive packages, any one of which could be invoked for carrying out appropriate deductions. A full system of the kind we have been proposing is still far in the future, but even limited working systems can be expected to be useful to modelers.

DATE
ILME